A feedback approach and its learning algorithmfor over complete blind source separation

ISPACS 2006

2006 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS2006)
Yonago Convention Center, Tottori, Japan

A Feedback Approach and Its Learning Algorithm for Over Complete Blind Source Separation

Kenji Nakayama Haruo Katou Akihiro Hirano
Graduate School of Natural Science and Technology, Kanazawa Univ.
Kakuma-machi, Kanazawa, 920-1192, Japan
Tel: +81-76-234-4896, Fax:+81-76-234-4900
E-mail:nakayama@t.kanazawa-u.ac.jp

***Abstract*— In blind source separation (BSS) applications, the number of the signal sources is not known. When the number of the sensors is less than that of the signal sources, this problem is called 'Over Complete BSS' (OC-BSS), which is a difficult problem due to lack of information in observations.**

In this paper, a feedback approach and its learning algorithm are proposed for the OC-BSS. The number of the outputs of an umixing block is set to be equal to that of the sensors. By assuming some condition, at least one output can separate a single signal source. This output is fed back to the inputs of the unmixing block, and is subtracted from the observations, in order to reduce the number of equivalent signal sources. Two kinds of feedback methods are proposed. One of them is direct subtraction and the other is sample elimination based on histogram of the feedback signal and the observed signals. The modified observations are further separated. The same process is repeated until all signal sources are separated.

Performance of the proposed method is evaluated through computer simulation. The proposed method can improve a signal to interference ratio by the several dB compared to the conventional methods.

I. Introduction

Signal processing, including noise cancellation, echo cancellation, equalization of transmission lines, estimation and restoration of signals have become a very important research area. Also, separation of mixed signals, for instance, group talking, is required in many situations. In some cases, however, we do not have enough information about signals and their interference. Furthermore, their mixing and transmission processes are not well known in advance. In these situations, blind source separation (BSS) technology using statistical properties of signal sources have become very important [1],[2],[4],[6].

In many real applications, the number of the signal sources cannot be estimated. The number of the sensors is usually different from that of the signal sources. When the number of the sensors is less than that of the signal sources, this problem is called 'Over Complete' BSS (OC-BSS). The OC-BSS is a difficult problem, due to lack of information in the observations about the signal sources. Therefore, the OC-BSS requires another information concerning the signal sources, besides the observed signals.

Several kinds of conventional methods have been proposed, which mainly use the histogram of the observed signals as the additional information [9],[10],[11]. However, separation performance is not well for practical applications.

In this paper, a new network structure and its learning algorithm are proposed. Each signal source is separated one by one in a feedback structure. The histogram of the observed and separated signals are also used as the additional information.

II. Feedback Approach to Over Complete BSS

For simplicity, 3 signal sources and 2 sensors are used. A block diagram is shown in Fig.1.

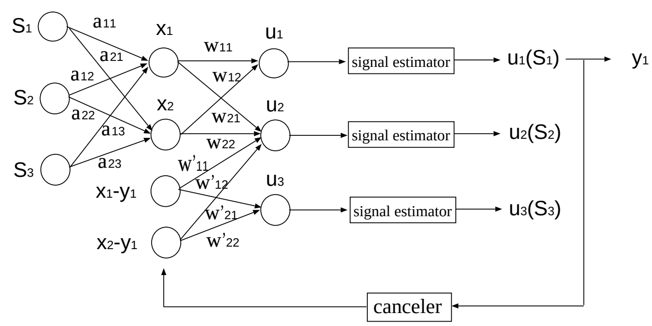

Fig. 1. Feedback network for over complete BSS with 3 signal sources and 2 sensors.

Letting N be the number of the signal sources, the number of the sensors M is set to be $M \geq [N/2+1]$, where $[X]$ means an integer number not exceed X. Under this condition, at least one output can include a single signal source. Because learning algorithms of the BSS make the outputs of the unmixing block to be statistically independent. Therefore, one signal source can be separated. On the other hand, a single output can include a plural number of the signal sources. For example, u_1 includes s_1 and u_2 includes s_2 and s_3.

Assume s_1 is separated in u_1. Since u_1 includes only a single voice, it is selected as the final output y_1. y_1 is fed back to the inputs of the unmixing block, and subtracted from the input signals x_1 and x_2, in order to eliminate the s_1 component in x_1 and x_2. Let the resulting x_1 and x_2 be x_1' and x_2', respectively. x_1' and x_2', which include only s_2 and s_3, are separated through another unmixing block represented with w_{ji}'. In this case, the number of the sensors and the outputs are the same as that of the signal sources. Then, they can be separated by the conventional method.

III. Signal Source Separation in First Phase

A. Theoretical Analysis of Source Separation

The network shown in Fig.1 is taken into account here. Furthermore, the mixing process is assumed to be an instantaneous process, that is a_{ji} do not include any time delay. The signal sources, the mixing block, the observed signals and the outputs of the unmixing block are related by

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{2}$$

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \tag{3}$$

Furthermore, the above equations are expressed by using vectors and matrices as follows:

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s} \tag{4}$$

$$\boldsymbol{u} = \boldsymbol{W}\boldsymbol{x} = \boldsymbol{W}\boldsymbol{A}\boldsymbol{s} = \boldsymbol{H}\boldsymbol{s} \tag{5}$$

Assume s_1 is separated in u_1, and s_2 and s_3 are separated in u_2. A signal to interference ratio in the 1st and the 2nd outputs, $u_1(n)$ and $u_2(n)$, are evaluated by

$$SIR_1 = 10\log_{10}\left(\frac{h_{11}^2}{h_{12}^2 + h_{13}^2}\right) \quad \text{[dB]} \tag{6}$$

$$SIR_2 = 10\log_{10}\left(\frac{h_{22}^2 + h_{23}^2}{h_{21}^2}\right) \quad \text{[dB]} \tag{7}$$

In order to maximize SIR_1 and SIR_2, the following conditions should be satisfied.

$$h_{11}^2, h_{22}^2, h_{23}^2 \quad \text{to be constant.} \tag{8}$$

$$h_{12}^2, h_{13}^2, h_{21}^2 \quad \text{to be minimized.} \tag{9}$$

The weights w_{kj} of the unmixing block can be adjusted so as to satisfy the above conditions resulting in theoretical performance. Furthermore, a balance of the s_1 and s_2 components is also important, which is evaluated by

$$R = \frac{|h_{23}|}{|h_{22}|} \tag{10}$$

B. Learning Algorithm

Conventional learning algorithms can be basically applied to the group separation, that is separating s_1 and (s_2, s_3). The learning algorithm using a mutual information as a cost function, and adjusts the weights following the natural gradient method [7],[8] is applied to this problem.

$$l(\boldsymbol{W}) = -\log|\det(\boldsymbol{W})| - \sum_{k=1}^{M}\log p_k(u_k) \tag{11}$$

$$\boldsymbol{W}(n+1) = \boldsymbol{W}(n) + \eta[\boldsymbol{\Lambda}(n) - <\phi(\boldsymbol{u}(n))\boldsymbol{u}^T(n)>]\boldsymbol{W}(n) \tag{12}$$

The operation $<>$ is time averaging. p_k is a probability density function of u_k. In order to stabilize a learning process, ϕ must satisfy Eq.(13), where p' is a 1st derivative of p, which is also a probability density function [7]. Several methods have been proposed for this purpose [3],[5]. In this paper, ϕ is controlled by Eq.(14), where κ_4 is kurtosis.

$$\phi(\boldsymbol{u}(n)) = \frac{p'(\boldsymbol{u}(n))}{p(\boldsymbol{u}(n))} \tag{13}$$

$$\phi(\boldsymbol{u}(n)) = a\tanh(\boldsymbol{u}(n)) + (1-a)\boldsymbol{u}^3(n) \tag{14}$$

$$a = \frac{1-\exp(-2.1\kappa_4 - 2.5)}{1+\exp(-2.1\kappa_4 - 2.5)} \tag{15}$$

C. Learning Control by Histogram of Observations

A estimation method for the mixed process by using histogram of the observations has been proposed [11]. This approach is taken into the learning process in this paper. The observation $\boldsymbol{x}(n)$ is projected onto a hyper cube, resulting in $\boldsymbol{v}(n) = \pi(\boldsymbol{x}(n))$. Example of the histogram is shown in Fig.2.

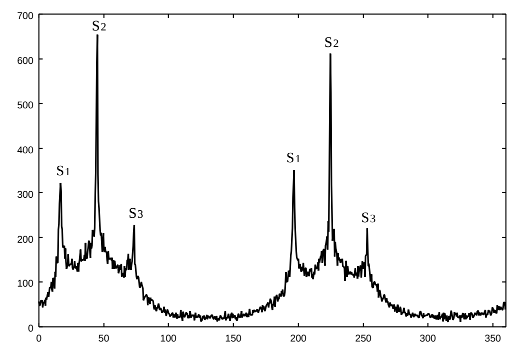

Fig. 2. Example of histogram of observations projected onto hyper cube. Horizontal axis indicates angle and vertical axis is histogram.

The learning algorithm is modified as follows:

$$\boldsymbol{w}^*(n) = \arg\max_{\boldsymbol{w}_j(n)}\{\boldsymbol{v}^T(n)\boldsymbol{w}_j(n)\} \tag{16}$$

$$\boldsymbol{w}^*(n+1) = \boldsymbol{w}^*(n) - \eta[<\phi(\boldsymbol{v}(n))\boldsymbol{v}^T(n)>]\boldsymbol{w}^*(n) \tag{17}$$

Idea behind the above learning algorithm is to update the weight vector, which is most close to the observation vector. In Eqs.(16) and (17), the norm of the weights $\boldsymbol{w}_j(n)$, $\boldsymbol{w}^*(n)$ and $\boldsymbol{w}^*(n+1)$ are normalized to be unity.

D. Simulation and Discussions

Two male voices and one female voice are used as the signal sources. The mixing process is determined as follows: $a_{11} = a_{23} = 1$, $a_{13} = a_{21} = 0.3$, $a_{12} + a_{22} = 1.4$. Furthermore, a ratio of a_{12} and a_{22} is defined by

$$\alpha = \frac{a_{12}}{a_{22}}, \qquad 0 < \alpha \leq 1 \tag{18}$$

When $\alpha = 1$, s_2 locates at the middle point between two sources, then separating s_2 into $u_2(n)$ is very difficult. On the other hand, when α takes a small value, s_2 locates close to s_3, then s_2 and s_3 are easily separated in $u_2(n)$.

Simulation results are shown in Figs.3, 4 and 5. 'Conventional method' means the learning process defined by Eq.(12).

'Proposed method' means the learning process defined by Eqs.(14) through (17). SIR_1 of the proposed method is almost the same as the theoretical curve, and is better than the conventional. SIR_2, whose theoretical value is infinite, is more than 16dB. The balance R is ideally unity. However, the theoretical value cannot be unity, and some imbalance cannot be avoided. R by the proposed method is slightly increased for a large α.

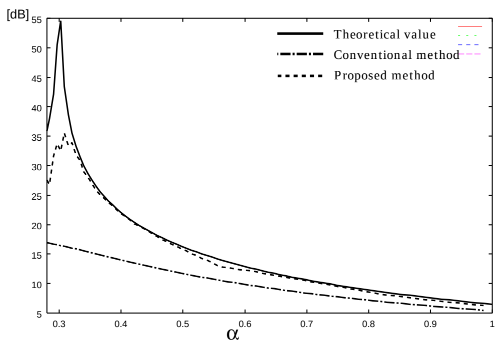

Fig. 3. SIR_1 defined by Eq.(6) in 3-2-2 BSS by using voice signals.

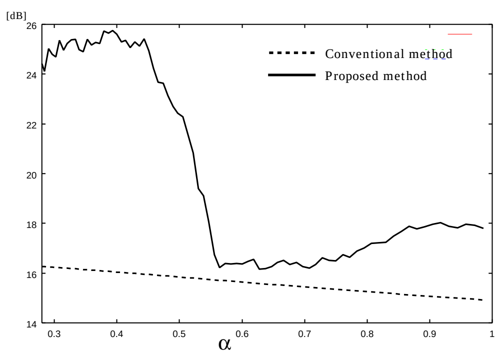

Fig. 4. SIR_2 defined by Eq.(7) in 3-2-2 BSS by using voice signals.

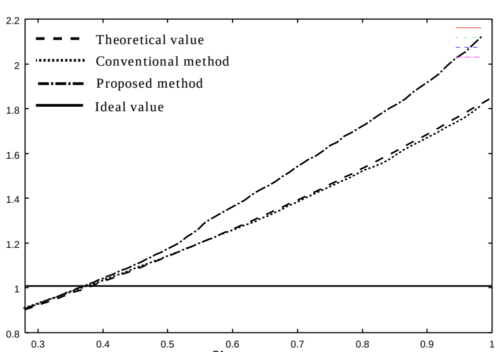

Fig. 5. R defined by Eq.(10) in 3-2-2 BSS by using voice signals.

IV. Elimination of A Single Voice Through Feedback

A. Estimation of Mixing Block

Basically speaking, the mixing block is unknown. However, it can be estimated by using histogram of the observations [11].

Figure 6 shows distribution of the observations $\boldsymbol{x}$. The horizontal and the vertical axes indicate $x_1(n)$ and $x_2(n)$, respectively. $\hat{\boldsymbol{a}}_i(n)$ is the estimation of $\boldsymbol{a}_i = [a_{1i}, a_{2i}]^T$, which follows the distribution of $\boldsymbol{x}_i(n)$.

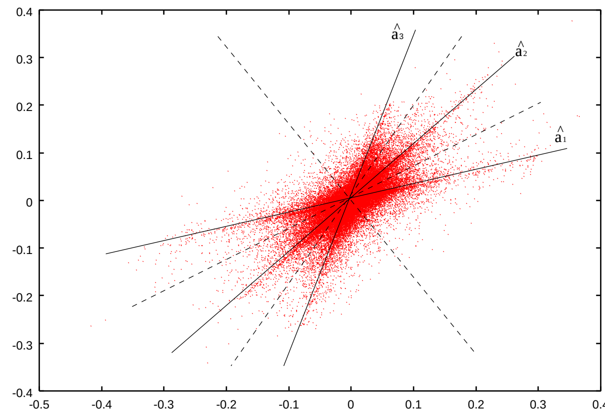

Fig. 6. Distribution of $x_1(n)$ and $x_2(n)$, which are expressed by the horizontal axis and the vertical axis, respectively. $\alpha = 0.428$.

B. Elimination of A Single Voice Through Two Methods

1) Direct Elimination: Suppose $u_1(n)$ includes a single voice, that is $s_1(n)$, and the mixing process is estimated. $u_1(n)$ is selected as the final output $y_1(n)$ as shown in Fig.1. $x_1(n)$ and $y_1(n)$ are expressed by

$$x_1(n) = a_{11}s_1(n) + a_{12}s_2(n) + a_{13}s_3(n) \qquad (19)$$

$$y_1(n) = (a_{11}w_{11} + a_{21}w_{12})s_1(n) \qquad (20)$$

Following these relations, $y_1(n)$ is subtracted from $x_1(n)$ as:

$$x_1'(n) = x_1(n) - \frac{a_{11}y_1(n)}{a_{11}w_{11} + a_{21}w_{12}} \qquad (21)$$

$x_1'(n)$ does not include $s_1(n)$.

2) Elimination based on Histogram: $s_1(n)$ included in $y_1(n)$ is subtracted from $x_2(n)$ based on the histogram of $y_1(n)$ and $x_2(n)$. $y_1(n)$ and $x_2(n)$ are projected onto the hyper cube, and their histogram are obtained. Samples of the observations following the histogram of $y_1(n)$ are randomly selected and are set to be zero. After that, $x_2(n)$ is denoted $x_2'(n)$. The histogram of $x_2'(n)$ will approach to that of $s_2(n)$ and $s_3(n)$. At the same time, it is controlled so as to follow the super Gaussian distribution.

C. Source Separation in Second Phase

After $x_1(n)$ and $x_2(n)$ are modified to $x_1'(n)$ and $x_2'(n)$, they are separated following the proposed learning algorithm. For instance, $s_2(n)$ is separated in $u_2(n)$, and $s_3(n)$ is separated in $u_3(n)$. In the second phase, the weights from $x_1'(n)$ and $x_2'(n)$ to $u_2(n)$ and $u_3(n)$ are adjusted.

D. Simulation and Discussion

The signal to interference ratio is evaluated by

$$SIR'_i = 10\log_{10}\left(\frac{\sum s_i^2(n)}{\sum(s_i(n)-y_i(n))^2}\right) \quad \text{[dB]} \qquad (22)$$

The power of $s_i(n)$ and $y_i(n)$ are normalized to be the same.

First, the histogram of $x'_2(n)$ is shown in Fig.7. Compared to the histogram of the observations shown in Fig.2, the histogram of $s_1(n)$ is eliminated.

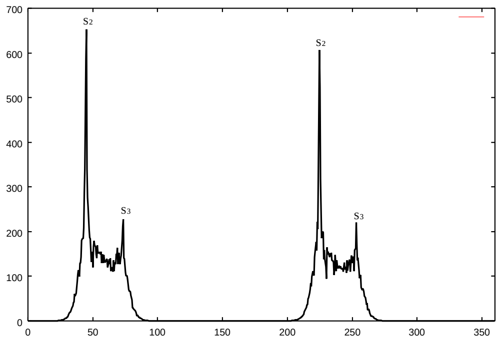

Fig. 7. Histgram of $x'_2(n)$, which is obtained by eliminating histogram of $y_1(n)$, that is $s_1(n)$, in feedback process.

The signal to interference ratios are shown in Figs.8, 9 and 10. 'Conventional Method' means the Shortest-Path method based on the histogram of the observations proposed by [11].

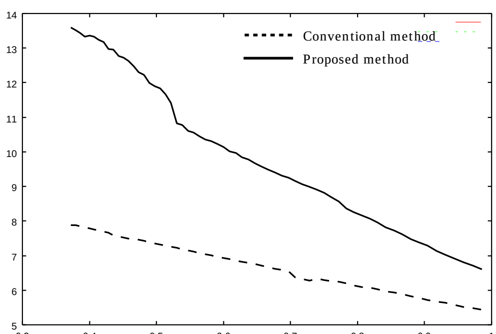

Fig. 8. SIR'_1 of 3-2-2 BSS. Horizontal axis is α and vertical axis is SIR'_1 in dB.

SIR'_1 decreases as α increases. In the first phase, since $s_1(n)$ and $s_2(n)+s_3(n)$ are separated, then a small α is better for separation. On the contrary, SIR'_2 and SIR'_3 increase as α increases. Because, in the second phase, $s_2(n)$ and $s_3(n)$ are separated. As α increases, $s_2(n)$ and $s_3(n)$ locate far from each the other. This means separation of $s_2(n)$ and $s_3(n)$ becomes more easy. The proposed method is better than the conventional method by $1 \sim 3$dB.

V. Conclusions

A feedback approach has been proposed for the over complete BSS. One output separates a single signal source, which is is fed back and subtracted from the observations in order to reduce the equivalent number of the signal sources. Two kinds of subtraction methods have been proposed. The signal to interference ratio can be improved by $1 \sim 3$dB.

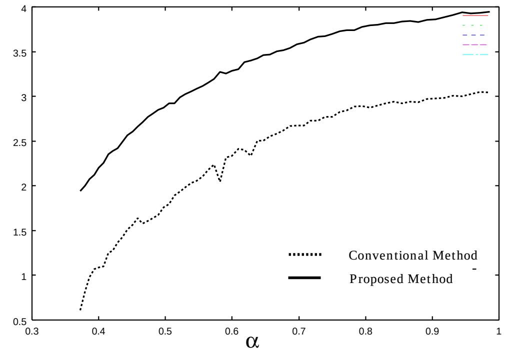

Fig. 9. SIR'_2 of 3-2-2 BSS. Horizontal axis is α and vertical axis is SIR'_2 in dB.

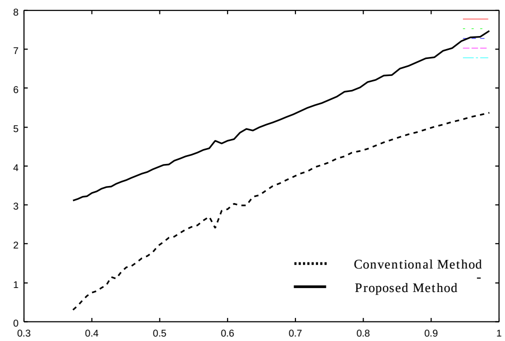

Fig. 10. SIR'_3 of 3-2-2 BSS. Horizontal axis is α and vertical axis is SIR'_3 in dB.

References

[1] C.Jutten and Jeanny Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture," Signal Processing, 24, pp.1-10, 1991.

[2] H.L.Nguyen Thi and C.Jutten, "Blind source separation for convolutive mixtures," Signal Processing, vol.45, no.2, pp.209–229, March 1995.

[3] K.Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by kurtosis for blind source separation," IEEE INNS, Proc. IJCNN'2002, Honolulu, pp.1234-1239, May 2002.

[4] K.Nakayama, A.Hirano and A.Horita, "A learning algorithm with adaptive exponential stepsize for blind source separation of convolutive mixtures with reverberations," IEEE INNS, Proc. IJCNN'2003 July 2003.

[5] K. Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by estimated output pdf for blind source separation," Proc. ICA2003, Nara, Japan, pp.427-432, April 2003.

[6] A.Horita, K.Nakayama, A.Hirano and Y.Dejima, "Analysis of signal separation and signal distortion in feedforward and feedback blind source separation based on source spectra," IJCNN'05, Montreal, Aug. 2005.

[7] S. Amari, T. Chen and A. Cichocki, "Stability analysis of learning algorithms for blind source separation", Neural Networks, vol.10, no8, pp.1345-1351, 1997.

[8] S.Amari, "Natural gradient works efficiently in learning," Neural Computation, vol.10, no2, pp.251-275, Feb. 1998.

[9] K.Waheed and F.M.Salem, "Algebraic overcomplete independent component analysis," ICA2003, Nara, Japan, April 2003.

[10] M.S.Lewicki and T.J.Sejnowski, "Learning overcomplete representations," Neural Computation, 2001.

[11] A.Jung, F.J.Theis, C.G.Puntonet and E.W.Lang, "FastGeo - A histogram based approach to linear geometric ICA," Proc. of ICA2001, pp.418-423, April 2001.

[12] A.D.Cheveigne and H.Kawahara, "YIN, A fundamental frequency estimator for speech and music," J.Acoust. Soc. Am., 2001.